I would like to thank the authors for their reply to my initial review.

The first major point of my initial review regarding molybdenum converters was reasonably addressed. The second major point was addressed, but I still have an open comment regarding the discussion of AMFs. Otherwise, I see the scientific content of the manuscript fit for publication. I have attached a fairly long list of fairly basic comments. I recommend publication after their consideration:

- l. 24: This sounds like WRF-Chem has issues with photo-chemistry under cloudy conditions, but in l. 359-364, the authors argue that the modelled jNO_2 agrees to satellite measurements. It would be good to express more clearly, what other aspect of photo-chemistry WRF-Chem could be struggling with.

- l. 29-37: This leaves out a few considerable sources, e.g. lightning NO_X and soil emissions.

- l. 53: column amounts → column densities

- l. 59: The sentence about oversampling is not clear (neither grammatically, nor semantically). To my knowledge, downscaling means to achieve truly higher resolution, not just a higher „nominal“ resolution by cutting a pixel into *N* parts with the same value. It is also not clear what purpose downscaling/oversampling has in the context of the paper.

- l. 63: This is incorrect: The measurements are of *spectra*, or perhaps, in the broader sense, of *column densities*. The *tropospheric column densities* require additional retrieval steps, that go far beyond pure measurement.

- l. 80: oxidation environmen*t* → oxidative capacity?

- l. 88: Even without dry deposition, a more shallow boundary layer will enhance the surface concentrations of trace gases, simply due to compression of the air mass into a smaller total volume.

- l. 99-102: It is not clear what the difference between point 1) and point 2) is, except that 1) refers to TROPOMI and 2) to geostationary satellites. Besides, the two points seem to express the same, namely to quantify the cloud-bias of NO_2 satellite measurements.

- l. 108-109: I suggest to make it more clear that the intent of the measurement is to quantify NO_2 and NO, and that the conversion of NO_y is an unwanted side-effect.

- l. 111: NO_2* (meaning „true NO_2 + false NO_2“) is not explicitly defined.

- Equation (1): The formula of Lamsal et al. (2008) contains an additional term for alkyl nitrates that does not occur here. This is fine, as WRF-Chem simulations often do not include alkyl nitrates in the first place, but it should be mentioned.

- l. 132-133: Does this statement refer only to TROPOMI or also to TEMPO?

- l. 244: Redundant „13:30“

- l. 248: The two sentences with „17.2 %“ appear to state the same.

- l. 260: Here (and in other places), change „no cloud days“ to „cloud free days“.

- Fig. 3: Mention in the caption that one scatter point corresponds to one measurement station.

- l. 271: decreased photolysis is also another reason.

- l. 286-287: „small (but larger)“ is not clear to me.

- l. 314-326: A few (minor) comments to this section: Small AMFs are not caused by the model „filling in“ the missing NO_2, but the other way around: If the AMFs are small, it implies that the satellite had reduced sensitivity, and the retrieval depends more on a priori assumptions (here: profile shapes taken from the 1° × 1° model TM5). Note, that clouds can also have an opposite

effect, where the sensitivity can be *enhanced* above bright clouds.

In its current form the section generally goes in the right direction, as it identifies the challenge of computing AMFs under cloudy conditions. However it still does not explicate the fundamental limitation I mentioned in my first review. I try to explain my concern again:

Assume, that the TM5 model works fine under cloud-free conditions, but has shortcomings under cloudy conditions. This is theoretically possible and cannot be ruled out (the authors' own results described in section 3.4 seem to be affected by this). For example, if the model had biased estimates of radiative transfer through clouds, this would result in faulty NO_2 photolysis rates, and possibly faulty NO_2 profile shapes. Note, that there exist plenty of other physical effects that could be modelled incorrectly. Then, the comparison is made between cloudy days (that are affected by said bias) and cloud-free days (that are unaffected by said bias). The differences one sees in this comparison could (in theory) be produced by the faulty a priori assumptions instead of actual differences in the NO_2 abundance. The authors mention a very similar notion wrt. the surface NO_2 in their WRF-Chem simulation in l. 356-357. The same logic should apply to vertical column density retrievals.

In other words, the presented work can answer the question: „How different are *retrieved* NO_2 VCDs under cloudy/cloud-free conditions?" but not „How different are the *actual* NO_2 VCDs under cloudy/cloud-free conditions?". The first question acknowledges that our a priori knowledge could be faulty, while the second question asks for more fundamental results that cannot be obtained in the presented methodological framework. I think the authors must distinguish clearly between pure measurement and retrieval, also wrt. to specific phrasing, e.g. in l. 63, see above.

- l. 346: Two comments here: Firstly, as stated by the authors previously, NO_2^* is also affected by PAN and alkyl nitrates. Secondly, the issue with the molybdenum-based measurements is not that HNO_3 is a terminal NO_2 sink, but that it simply introduces a strong measurement bias (as in: the measurements are not correct).